UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/734,228	12/11/2000	Helmut Lucke	450117-02965	5435

20999	7590	09/28/2006

FROMMER LAWRENCE & HAUG
745 FIFTH AVENUE- 10TH FL.
NEW YORK, NY 10151

EXAMINER
JACKSON, JAKIEDA R

ART UNIT	PAPER NUMBER
2626	

DATE MAILED: 09/28/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 10/03)

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
  Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) ☒ Responsive to communication(s) filed on *20 July 2006*.

2a) ☒ This action is **FINAL**.   2b) ☐ This action is non-final.

3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) ☒ Claim(s) *1,2,4,5,9-12 and 14-21* is/are pending in the application.

    4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5) ☐ Claim(s) _____ is/are allowed.

6) ☒ Claim(s) *1-2, 4-5, 9-12 and 14-21* is/are rejected.

7) ☐ Claim(s) _____ is/are objected to.

8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) ☐ The specification is objected to by the Examiner.

10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.

    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

    a) ☐ All   b) ☐ Some * c) ☐ None of:

      1. ☐ Certified copies of the priority documents have been received.

      2. ☐ Certified copies of the priority documents have been received in Application No. _____.

      3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

    * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☐ Information Disclosure Statement(s) (PTO/SB/08)
    Paper No(s)/Mail Date _____.

4) ☐ Interview Summary (PTO-413)
    Paper No(s)/Mail Date. _____.

5) ☐ Notice of Informal Patent Application

6) ☐ Other: _____.

DETAILED ACTION

Response to Amendment

1.   In response to the Office Action mailed April 21, 2006, applicant submitted an

amendment filed on July 20, 2006, in which the applicant traversed and requested

reconsideration with respect to **claim 1**.

Response to Arguments

2.   Applicants argue that nothing has been found in Jiang, Ehsani, Kimura or Chou

taken alone o in combination would teach or suggest said pragmatic information

includes connecting information connecting said sub-phrases to actual situation,

application and/or action. Additionally, the cited combinations do not teach or suggest

that a language model is used containing at least a recognition grammar built up by at

least a low-perplexity part and a high-perplexity part, each of which being representative

for distinct low and high perplexity classes of speech elements. Finally, the cited

combinations do not teach or suggest that word classes are used as classes for speech

elements or fragments, as recited in claim 1.

Ehsani teaches that the operation of a voice-interactive application entails

processing acoustic, syntactic, semantic and ***pragmatic information*** derived from the

user input in such a way as to ***generate a desired response*** from the application

(column 11, paragraph 0216). Ehsani also teaches that if n-gram is part of a larger

string collocation the choice of words adjacent to the phrase boundary will be very

small, because of the internal constraint of the collocation. Conversely, the likelihood

that a particular word will follow is very high. For example, the word following the

trigram "to a large" will almost always be "extent" which means the perplexity is low, and

the trigram is subsumed under the fixed collocation "to a large extent." On the other

hand, a large number of different words can precede or follow the phrase "to a large

extent", and the probability that any particular word will follow is very small (close to 0),

columns 5-6, paragraph 0102.

Further Chou teaches a language model that contains at least a recognition

grammar (column 6, line 66 – column 7, line 5) built up by at least a low-perplexity part

and a high-perplexity part, each of which being representative for distinct low-and high-

perplexity classes of speech elements (column 2, lines 61-65) and that word classes are

used as classes for speech elements or fragments (column 8, lines 18-32). This leads

to a robust understanding of the utterance (column 5, lines 27-49).

Therefore, applicant's arguments are not persuasive.

Claim Rejections - 35 USC § 103

3.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set
> forth in section 102 of this title, if the differences between the subject matter sought to be patented and
> the prior art are such that the subject matter as a whole would have been obvious at the time the
> invention was made to a person having ordinary skill in the art to which said subject matter pertains.
> Patentability shall not be negatived by the manner in which the invention was made.

4.   **Claims 1-2, 4-5, 9-12 and 14-21** are rejected under 35 U.S.C. 103(a) as being

unpatentable over Jiang et al. (U.S. Patent No. 6,539,353), hereinafter referenced as

Jiang in view of Ehsani et al. (U.S. Publication No. 2002/0128821), hereinafter

referenced as Ehsani and in further view of Kimura et al. (USPN 6,067,510), hereinafter

referenced as Kimura and in further view of of Chou et al. (U.S. Patent No. 5,797,123),

hereinafter referenced as Chou.

    As per **claims 1 and 14**, Jiang discloses a method and apparatus for recognizing

speech, comprising:

    (a) the steps of receiving a speech phrase (100, FIG. 2);

    (b) generating a signal being representative to said speech phrase using A/D

converter (102, FIG.2);

    (c) using feature extractor for pre-processing and storing said signal (104, FIG.

2);

    (d) generating from said pre-processed signal at least one series of hypothesis

speech elements (Col. 1, line 51-53);

    (e) determining at least one series of words being most probable to correspond

to said speech phrase by applying a predefined language model to at least said series

of hypothesis speech elements (Col. 4, lines 13-16),

    wherein the step of determining said series of words further comprises the steps

of:

    (1) identifying a hypothesis string consisting of sub-word units (Col. 1, lines 52-

55) then continuing determining words or combinations of words and which are

consistent with said seed sub-phrase as at least a first successive sub-phrase which is

contained in said received speech phrase (Col. 6, lines 38-46 with Col. 5, lines 28-51

and Col. 4, lines 33-44), but lacks identifying and extracting word classes of high-

perplexity, applying a compiler, merging the sub-word-unit grammars with the remaining

low-perplexity part and inserting additional information.

Ehsani discloses phrase-based dialogue modeling method for producing a low-

perplexity recognition grammar from a conventional grammar having semantic

information including a description between sub-phrases (column 3, paragraphs 0034-

0043) comprising:

(a) identifying and extracting word classes (trigram subsumed under the fixed

collocation) of high-perplexity (very high perplexity) from the conventional grammar

(column 5, paragraphs 0100-0102);

(b) generating a phonetic, phonemic and/or syllabic description (phone models

and phonetic dictionary; column 11, paragraph 0217) of high-perplexity word classes

(very high perplexity), in particular by applying a sub-word-unit grammar compiler to

them (column 11, paragraphs 0211-0214 with column 10, paragraphs 0199-0200), to

produce a sub-word-unit grammar for each high-perplexity word class (column 5,

paragraphs 0100-0102);

(c) merging sub-word-unit grammars (combining) with remaining low-perplexity

part of the conventional grammar to yield said low-perplexity recognition grammar

(column 4, paragraphs 0064 with column 6, paragraph 0107), to measure the strength

of certain collocations;

wherein said seed sub-phrase is recognized with an appropriate high degree of

reliability, such that segments of speech that are recognized with high reliability are

used to constrain the search in other areas of the speech signal where the language

model employed cannot adequately restrict the search (column 3, paragraph 0059,

column 5, paragraph 100 and column 11, paragraph 0221).

Therefore, it would have been obvious to one of ordinary skill in the art at the

time the invention was made to modify Jiang's method wherein it identifies and extracts

word classes of high-perplexity, applies a compiler, merges the sub-word-unit grammars

with the remaining low-perplexity part and constrain the search and provides pragmatic

information contained in a reliably recognizable part of the speech phrase that is useful

to explain another part of higher perplexity, to measure for determining the average

branching factor of a recognition network, for evaluating language models (column 5,

paragraph 0100) to generate a desired response from the application (column 11,

paragraph 0216).

Jiang in view of Ehsani discloses a method and apparatus for recognizing

speech, but does not specifically teach inserting additional information.

Kimura teaches inserting additional, higher order information (hierarchy),

including semantic (semantic features), between the sub-phrases, thereby decreasing

the burden of searching (greatly reduce labor and time to search; column 3, lines 43-

51), wherein the semantic information includes description of the sub-phrases (column

5, lines 38-56 with column 12, lines 22-26 and column 15, line 36-43).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Jiang in combination with Ehsani's method and apparatus such that it discloses inserting additional information, to sort and display words hierarchically in a particular order when displaying the words as candidates for substitution so that a time for retrieving the words can be reduced, as taught by Kimura (column 2, lines 1-6).

Jiang in view of Ehsani and Kimura does not disclose the use of low-perplexity and high-perplexity pads in the system.

Chou teaches a language model that contains at least a recognition grammar (column 6, line 66 – column 7, line 5) built up by at least a low-perplexity part and a high-perplexity part, each of which being representative for distinct low-and high-perplexity classes of speech elements (column 2, lines 61-65 with column 5, lines 27-67) and that word classes are used as classes for speech elements or fragments (column 8, lines 18-32). This leads to a robust understanding of the utterance (column 5, lines 27-49).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Jiang in combination with Ehsani and Kimura's method and apparatus, as taught in Chou, such that the subword-based speech recognizer is adapted to recognize a set of key-phrases using a set pf phrase sub-grammars which may advantageously be specific to the dialog. This may be useful in the sentence-level parsing and lead to a robust understanding of the utterance (column 5, lines 27-67).

As per **claim 2**, Jiang et al. disclose the use of a language model (110, FIG. 2) to provide additional information about the set of probabilities that a particular sequence of words will appear in the language of interest (Col. 4, lines 33-44)

As per **claims 4 and 5**, Jiang et al. discloses that language model (110, FIG- 2) is a compact trigram model that determines the probability of sequence of words based on the combined probabilities of three-word segment of the sequence. (Col.4, lines 41-44). Inherently, trigram language models take prepositional relationships of sub-phrases into account when calculating probabilities.

As per **claim 9**, Jiang et al. discloses the use of Hidden Markov Models for estimating probabilities for any sequence of sub-words generated by lexicon (Col. 4, lines 23-30).

As per **claim 10**, Jiang in view of Ehsani and Kimura does not disclose the insertion of high-perplexity word classes into hypothetic graph.

Chou teaches the insertion of functional words and filler phrases into the detection network to improve recognition of key-phrases (Col. 6, lines 47-56).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Jiang in view of Ehsani and Kimura's method and apparatus, as taught in Chou, in order to handle repeating speech patterns and thus speed up the search and improve recognition.

As per **claim 11**, Jiang in view of Ehsani and Kimura do not disclose the removal of candidates from the hypothetical graph.

Chou teaches the merging of the states of the key-phrase network, thus reducing

its size (Col. 7, lines 40-46).

Therefore, it would have been obvious to one of ordinary skill in the art at the

time the invention was made to modify Jiang combination with Ehsani and Kimura

method and apparatus, as taught in Chou, in order to prune the passed nodes while

doing the search through the hypothetical network and thus limit the possibility to

accidentally encroach upon the beginning of another phrase.

As per **claim 12,** Jiang in view of Ehsani and Kimura do not disclose restricting

the remaining part of the key-phrase.

Chou teaches placing additional constraints on the search that inhibit impossible

connections of key-phrases (Col. 6, lines 64-65).

Therefore, it would have been obvious to one of ordinary skill in the art at the

time the invention was made to modify Jiang in combination with Ehsani and Kimura

method and apparatus, as taught in Chou, in order to improve the speed of recognition

by quickly removing impossible combinations from the search graph and thus limiting

the search space.

As per **claim 15**, Jiang discloses a method and apparatus for recognizing

speech, but does not specifically include information relating to grammatical constraints

among said sub-seed.

Ehsani discloses a speech recognition method and apparatus including

information relating to grammatical constraints among said sub-seed (column 11,

paragraph 0221), to narrow down the hypotheses generated by the acoustic signal.

Therefore, it would have been obvious to one of ordinary skill in the art at the

time the invention was made to modify Jiang's method and apparatus wherein it

includes information relating to grammatical constraints among said sub-seed, to narrow

down the hypotheses generated by the acoustic signal, to come up with a number of

possible commands that are processed by the system (column 11, paragraph 0221).

As per **claim 16**, Jiang discloses a method and apparatus for recognizing

speech, but does not specifically include grammatical constraints for a name of a city.

Ehsani discloses a speech recognition method and apparatus including

grammatical constraints for a name of a city (column 10, paragraph 0196), to enable the

phrase thesaurus to be represented more compactly.

Therefore, it would have been obvious to one of ordinary skill in the art at the

time the invention was made to modify Jiang's method and apparatus wherein it

includes grammatical constraints for a name of a city, to enable the phrase thesaurus to

be represented more compactly thus decreasing the data storage capacity required to

store the data representing the phrase thesaurus (column 10, paragraph 0197).

As per **claim 17**, Jiang disclose a method and apparatus for recognizing speech,

but does not specifically discloses pragmatic information including digital postal code for

the city.

Ehsani teaches that the descriptors include businesses, restaurants, cities, etc.

(column 10, paragraph 0196), to enable the phrase thesaurus to be represented more

compactly.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Jiang's method and apparatus such that it includes a 5-digit postal code for the city, to allow the information to be received hierarchically with a large variety of different domains (column 2, paragraph 0022).

As per **claims 18 and 20**, Jiang disclose the method and apparatus for recognizing speech, but lacks wherein said seed sub-phrase recognized with an appropriate high degree of reliability is defined as a low perplexity part of said received speech phrase.

Ehsani disclose the method wherein said seed sub-phrase recognized with an appropriate high degree of reliability is defined as a low perplexity part of said received speech phrase (column 3, paragraphs 0034-0043 with column 4, paragraphs 0064 and column 6, paragraph 0107), to measure the strength of certain collocations. Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Jiang's method and apparatus wherein said seed sub-phrase recognized with an appropriate high degree of reliability is defined as a low perplexity part of said received speech phrase, as taught by Ehsani, to measure for determining the average branching factor of a recognition network, for evaluating language models (column 5, paragraph 0100).

As per **claims 19 and 21**, Jiang discloses the method wherein perplexity is defined as the complexity of the depth of search which has to be accomplished in conventional search graphs or search trees (column 4, lines 45-57).

Conclusion

5.   **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time

policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE

MONTHS from the mailing date of this action. In the event a first reply is filed within

TWO MONTHS of the mailing date of this final action and the advisory action is not

mailed until after the end of the THREE-MONTH shortened statutory period, then the

shortened statutory period will expire on the date the advisory action is mailed, and any

extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of

the advisory action. In no event, however, will the statutory period for reply expire later

than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the

examiner should be directed to Jakieda R. Jackson whose telephone number is

571.272.7619. The examiner can normally be reached on Monday through Friday from

7:30 a.m. to 5:00p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, David Hudspeth can be reached on 571.272.7843. The fax phone number

for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the

Patent Application Information Retrieval (PAIR) system. Status information for

published applications may be obtained from either Private PAIR or Public PAIR.

Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov. Should

you have questions on access to the Private PAIR system, contact the Electronic

Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a

USPTO Customer Service Representative or access to the automated information

system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

JRJ
September 25, 2006

DAVID HUDSPETH
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2600